Indicator: Cancer Mortality and Incidence (076 and 077)

Cancer is the second leading cause of death in the United States, accounting for about 23% of all deaths in 2001 (Anderson and Smith, 2003). The term "cancer" is used to characterize diseases in which abnormal cells divide without control. A cancerous cell loses its ability to regulate its own growth, control cell division, and communicate with other cells. Cancer cells can invade nearby tissues and can spread through the bloodstream and lymphatic system to other parts of the body (NCI, 2004). The risk of developing cancer increases with age and the environment (as broadly defined), genetic predisposition, certain viruses, and socioeconomic factors may all pay a role in the development and progression of the disease.

The contribution of environmental factors to the development of cancer has been and continues to be a major focus of research. Factors including individual food and beverage preferences, use of tobacco products, exposure to natural and medical radiation (including sunlight), workplace exposures, and pharmaceutical use as well as exposure to substances in the air, water and soil all may contribute individually (additive) or synergistically (i.e., an effect greater than the sum of each factor acting alone) to the development of cancer (NTP, 2004). Only in a small number of cases is it known what specific environmental factor(s) or condition(s) are responsible for the onset and development of cancers (NTP, 2004).

This indicator presents mortality and incidence for all U.S. cancers using data collected through the National Vital Statistics System, maintained by the National Center for Health Statistics (mortality) and the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) Program (incidence). The National Vital Statistics System (NVSS) registers virtually all deaths and births nationwide with data coverage from 1933 to present and from all 50 States and the District of Columbia. The SEER Program collects and publishes cancer incidence and survival data from 14 population-based cancer registries and three supplemental registries covering approximately 26 % of the U.S. population.

What the Data Show

In the United States, 553,768 and 557,271 people died of cancer in 2001 and 2002, respectively. Figure 076-1 presents the overall trends in cancer mortality for U.S. and the 10 EPA Regions for the time periods 1979-1998 and 1999-2001. The age-adjusted (2000 U.S. Standard Population) cancer mortality rate peaked around 1990 (216.0 per 100,000) and has been decreasing through 1998 with an overall rate of 200.7 per 100,000. A decline appears to continue between 1999-2001. In 1979, the Regional cancer mortality rates age-adjusted to the 2000 U.S. Standard Population ranged from 192.9 (Region 6) to 219.9 (Region 2) per 100,000 while age-adjusted rates in 1998 ranged from 177.3 (Region 8) to 210.4 (Region 3) per 100,000. Figure 076-2 presents a map of the 10 EPA Regions for the overall age-adjusted cancer mortality rates during 1999-2001. The rates on the map were divided into quartiles. The rates during that time period ranged between 172.7 (Region 8) and 211.3 (Region 3) per 100,000.

Like mortality, overall cancer incidence has been declining. Figure 077 presents the overall trend in cancer incidence between 1973-2000. Although a slow steady increase in cancer incidence occurred between 1973 and 1992 (peaked in 1992 with an age-adjusted cancer incidence of 509.9 per 100,000), overall incidence rates appear to have stabilized over the last ten years.

Differences exist in both mortality and incidence rates among gender, racial, and age groups, although rates within the different groups have been declining overtime. During 2001, those aged 65 and older had the highest cancer mortality (1,102.9 per 100,000) and incidence rates (2,158.7 per 100,000) compared to all other age categories. Both mortality and incidence rates are higher for men compared to women, and for blacks compared to whites. In 2001, the age-adjusted cancer mortality rate for black men was 330.9

per 100,000 compared to 239.2, 191.3, and 164.0 per 100,000 among white men, black women and white women, respectively (Anderson and Smith. 2003). Figure 077 presents age-adjusted cancer incidence for males and females between 1973-2000. Males peaked in 1992 with an age-adjusted incidence rate of 657.1 per 100,000 whereas the age-adjusted cancer incidence rates among women have slowly been increasing over time.

For cancer incidence, differences also occur between race and gender. The age-adjusted cancer incidence rate in 2001 for black men was 652.0 per 100,000 compared to 552.9 per 100,000 for white men; among women, the age-adjusted cancer incidence rate was 425.0 per 100,000 for white women compared to 380.1 per 100,000 among black women.

Indicator Limitations

- Cancer mortality rates are based on under-lying cause-of-death as entered on a death certificate by a physician. Some individuals may have had competing causes of death. "When more than one cause or condition is entered by the physician, the underlying cause is determined by the sequence of conditions on the certificate, provisions of the ICD, and associated selection rules and modifications" (CDC WONDER). Consequently, some misclassification of reported mortality might occur in individuals with competing causes of death.

- The difference in overall cancer mortality rate for 1999 quoted in ROE03 (201.6 per 100,000) versus the rate quoted in this updated reported (200.8 per 100,000) is that the rate in ROE03 was based on 1990 Census projections and the rate has since been adjusted for 2000 standard population

- The SEER Program currently only collects and publishes cancer incidence and survival data from 14 population-based cancer registries and three supplemental registries covering approximately 26 % of the U.S. population. However, the sampled population is considered representative of the national population (NCI, 2005).

- The International Classification of Diseases 9th Revision (ICD-9) codes are used to specify underlying cause of death for years 1979 - 1998. Beginning in 1999, cause of death is specified with the International Classification of Diseases 10th Revision (ICD-10) codes. The two revisions differ substantially, and to prevent confusion about the significance of any specific disease code, data queries are separate.

Data Sources

Cancer Mortality: CDC. CDC WONDER. Compressed Mortality File, Underlying Cause of Death. http://wonder.cdc.gov.
The complete web-link pathway from the CDC WONDER Home Page is:
— Mortality—underlying cause of death — Mortality for 1999–2001 with ICD 10 codes.
Note: ICD-10 codes C00–C97 are listed as malignant neoplasms

For 2002 data: National Center for Health Statistics (NCHS). 2004. Deaths: Final Data for 2002. National Vital Statistics Reports.53(5). October 12, 2004.
http://www.cdc.gov/nchs/data/nvsr/nvsr53/nvsr53_05.pdf

Cancer Incidence: National Cancer Institute (NCI). Surveillance, Epidemiology, and End Results (SEER)*Stat Database: Incidence –SEER 9 Regions Public Use, November 2003 Sub (1973-2001), NCI,

DCCPS, Surveillance Research Program, Cancer Statistics Branch, released April 15, 2004, based on the November 2003 submission. http://wonder.cdc.gov/seerJ.html.
The complete web-link pathway from the CDC WONDER Home Page (http://wonder.cdc.gov/) is:
— Cancer Incidence — SEER Cancer Query System — SEER Incidence and US Mortality Statistics
Selection criteria: Age Adjusted Rates, Standard Population = 2000 Census, Nine SEER Registry, All Sites; All Ages.

The raw numbers for each state were downloaded from the CDC WONDER mortality database (http://wonder.cdc.gov). The raw numbers for each state within a region were combined and age-adjusted rates (2000 U.S. Standard Population) were calculated.

References

Anderson and Smith. 2003. National Vital Statistics Report (NVSR). Deaths: Leading Causes for 2001. Volume 52, Number 9. November 7, 2003. http://www.cdc.gov/nchs/data/nvsr/nvsr52/nvsr52_09.pdf.

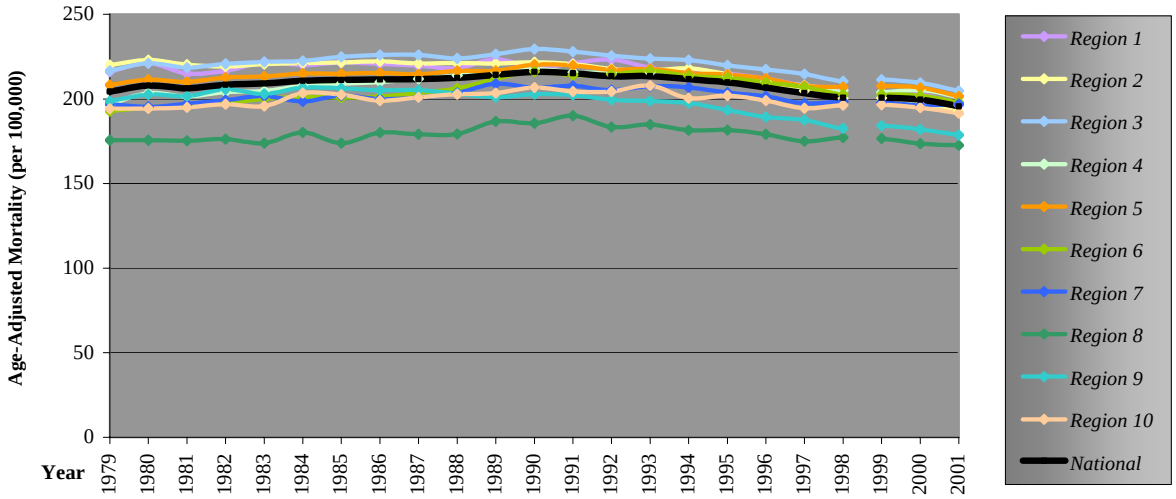
National Cancer Institute (NCI). Accessed October 7, 2004. http://cancer.gov/dictionary/

National Cancer Institute (NCI). Accessed January 11, 2005. http://seer.cancer.gov/about/

National Toxicology Program (NTP). 2004. Report on Carcinogens, Eleventh Edition; U.S. Department of Health and Human Services, Public Health Service, National Toxicology Program. Accessed February 2, 2005. http://ntp.niehs.nih.gov/ntp/roc/toc11.html

Graphics

Figure 076-1. Age Adjusted U.S. Cancer Mortality Rates, total U.S. and 10 EPA Regions, 1979-2001*



* - Due to differences in the ICD system used for classifying mortalities, data from 1979-1998 should not be directly compared to data from 1999-2001.
ICD-9 CODES: 140-208 (1979-1998); ICD-10 CODES: C00-C97 (1999-2001)

Figure 076-2. Age Adjusted U.S. Cancer Mortality Rates in Quartiles, 10 EPA Regions, 1999-2001

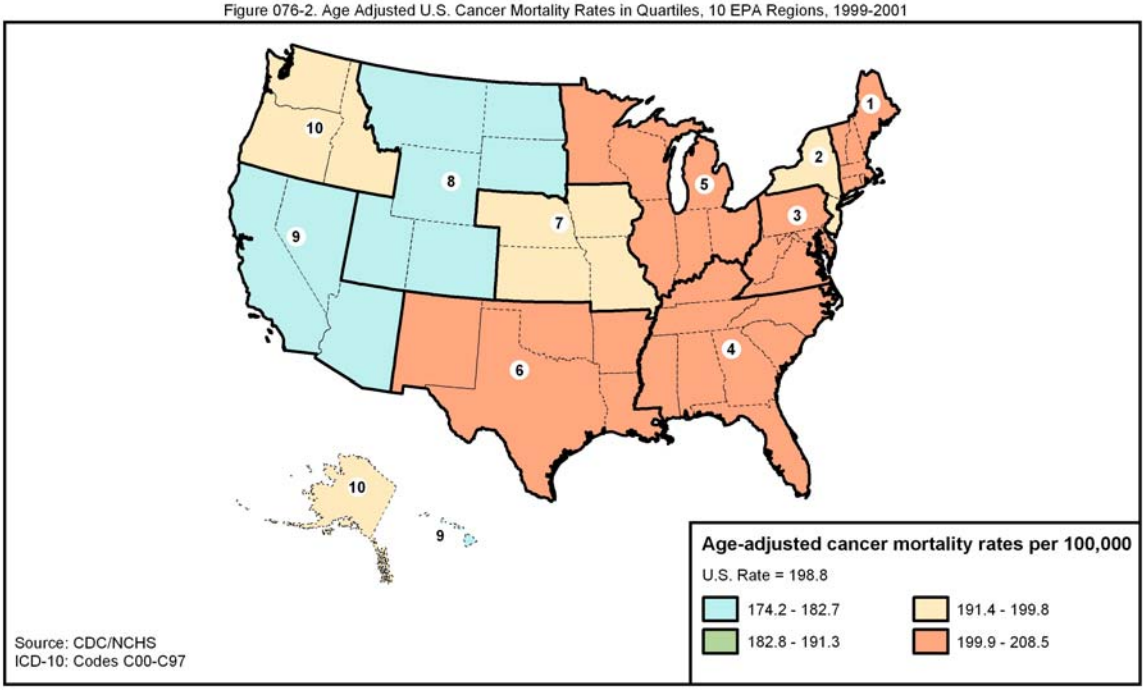
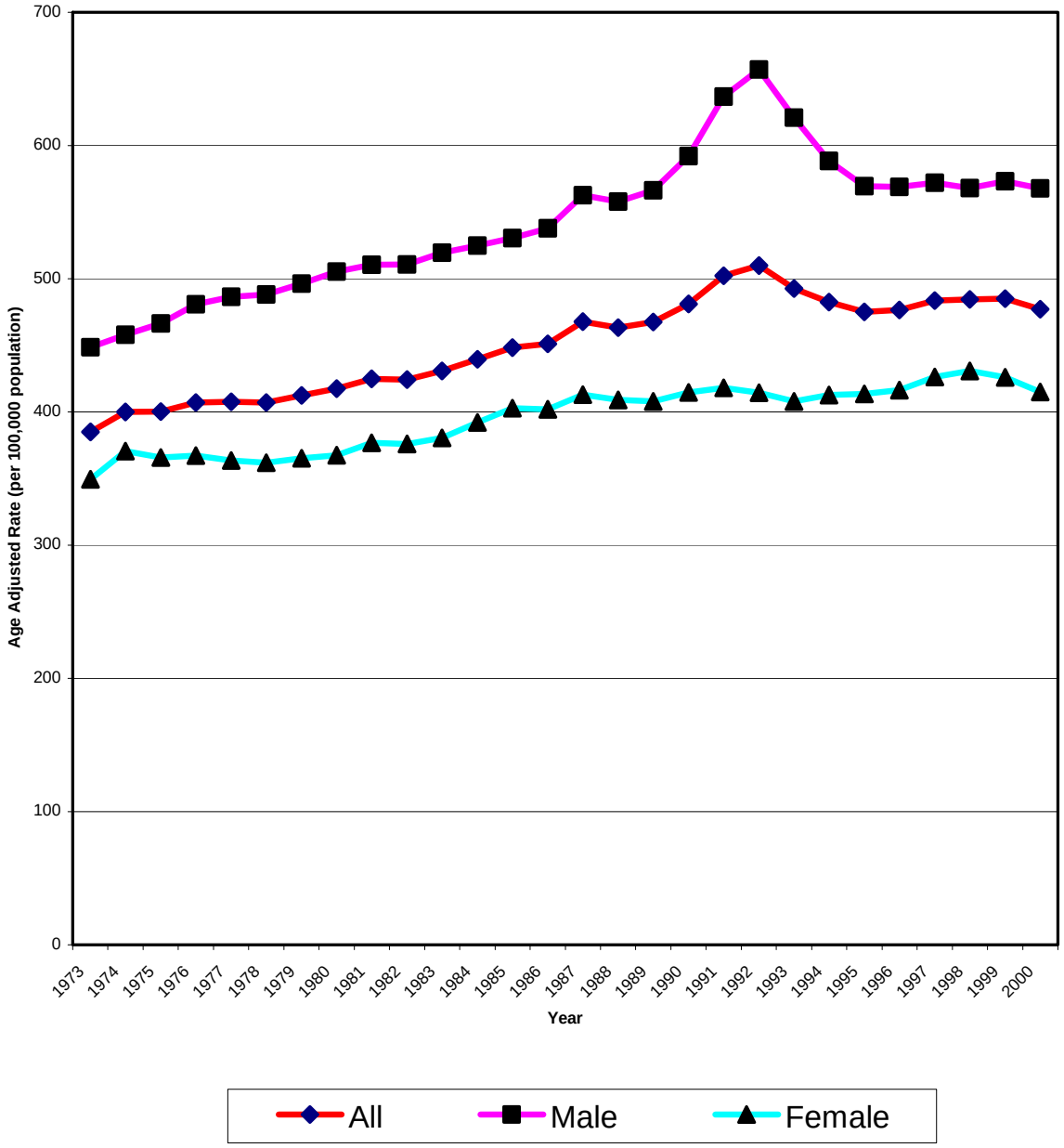


Figure 077: National Trends in Cancer Incidence Between 1973 and 2000



Source: NCI. Surveillance, Epidemiology, and End Results (SEER)*Stat Database:
Incidence - SEER 9 Regions Public Use, November 2003 Sub (1973-2001), NCI DCCPS,
Surveillance Research Program, Cancer Statistics Branch, Released April, 15, 2004.
http://wonder.cdc.gov/seerJ.html.

R.O.E. Indicator QA/QC

Data Set Name: CANCER INCIDENCE & MORTALITY
Indicator Number: 076 (89117)
Data Set Source: CDC, NCHS - mortality; NCI - incidence
Data Collection Date: ongoing
Data Collection Frequency: yearly
Data Set Description: Cancer Incidence & Mortality (combines indicators 076 & 077)
Primary ROE Question: What are the trends in human disease and conditions for which environmental pollutants are thought to be to risk factors including across population subgroups and geographic regions?

Question/Response

T1Q1 Are the physical, chemical, or biological measurements upon which this indicator is based widely accepted as scientifically and technically valid?

Yes. Mortality: The National Vital Statistics System (NVSS) is the oldest and most successful example of inter-governmental data sharing in Public Health and the shared relationships, standards, and procedures form the mechanism by which NCHS collects and disseminates the Nation's official vital statistics. The methodology for collecting vital statistics is standardized and outlined in "Model State Vital Statistics Act and Regulations" Revised April 1995, DHHS publication (PHS) 95-1115 (http://www.cdc.gov/nchs/data/misc/mvsact92aacc.pdf) Incidence: The Surveillance, Epidemiology, and End Results (SEER) Program of the National Cancer Institute is an authoritative source of information on cancer incidence and survival in the United States, and is considered the standard for quality among cancer registries around the world. The SEER Program currently collects and publishes cancer incidence and survival data from 14 population-based cancer registries and three supplemental registries covering approximately 26 percent of the US population. Information on more than 3 million in situ and invasive cancer cases is included in the SEER database, and approximately 170,000 new cases are added each year within the SEER coverage areas. An overview of SEER can be found at http://seer.cancer.gov/about/. Incidence: The Surveillance, Epidemiology, and End Results (SEER) Program of the National Cancer Institute is an authoritative source of information on cancer incidence and survival in the United States, and is considered the standard for quality among cancer registries around the world. The SEER Program currently collects and publishes cancer incidence and survival data from 14 population-based cancer registries and three supplemental registries covering approximately 26 percent of the US population. Information on more than 3 million in situ and invasive cancer cases is included in the SEER database, and approximately 170,000 new cases are added each year within the SEER coverage areas. An overview of SEER can be found at http://seer.cancer.gov/about/.

T1Q2 Is the sampling design and/or monitoring plan used to collect the data over time and space based on sound scientific principles?

Mortality: Yes. The National Vital Statistics System is responsible for the Nation's official vital statistics. These vital statistics are provided through State-operated registration systems. Standard forms for the collection of data and model procedures for the uniform registration of the events are developed and recommended for State use through cooperative activities of the States and the NCHS (http://www.cdc.gov/nchs/data/dvs/DEATH11-03final-ACC.pdf). U.S. Standard Death Certificates are revised periodically. Most state certificates conform closely in content and arrangement to the standard certificate recommended by NCHS and all certificates

contain a minimum data set specified by NCHS. Demographic information on the death certificate is provided by the funeral director based on information supplied by an informant. A physician, medical examiner, or coroner provides medical certification of cause of death. Incidence: Yes. The SEER Registries routinely collect data on patient demographics, primary tumor site, morphology, stage at diagnosis, first course of treatment, and follow-up for vital status. The SEER Program is the only comprehensive source of population-based information in the United States that includes stage of cancer at the time of diagnosis and survival rates within each stage. Additionally, studies are conducted on a yearly basis in SEER areas to evaluate the quality and completeness of the data being reported. (http://seer.cancer.gov/about/quality.html)

T1Q3 Is the conceptual model used to transform these measurements into an indicator widely accepted as a scientifically sound representation of the phenomenon it indicates?

Mortality: Yes. The data collected by NVSS are routinely referenced and used in epidemiological studies. Regional Mortality: Mortality rates age-adjusted for the 2000 U.S. standard population (rates per 100,000) for the years 1979 through 2001 were compiled through use of CDC WONDER. Maps were created for this same data. See attached file, Calculating data for GIS mortality maps and regional trends.doc , for more information on the calculation of rates. Data file for maps can be found in file MapDataSummaries.xls . Incidence: Yes. The data are published routinely referenced in scientific journals. Incidence: Yes. The data are published routinely referenced in scientific journals.

T2Q1 To what extent is the indicator sampling design and monitoring plan appropriate for answering the relevant question in the ROE?

Mortality: Virtually all deaths are registered with the NVSS nationwide. The temporal coverage of the data is from 1933 to present. Data are collected from all 50 States including the District of Columbia. Incidence: Collection began in 1973 and covers approximately 26 percent of the US population. However, this 26 percent is representative of the national population (http://seer.cancer.gov/registries/characteristics.html).

T2Q2 To what extent does the sampling design represent sensitive populations or ecosystems?

Mortality: The data set has nationwide death reporting, including sensitive populations. Incidence: This is nationally representative data

T2Q3 Are there established reference points, thresholds or ranges of values for this indicator that unambiguously reflect the state of the environment?

Not applicable

T3Q1 What documentation clearly and completely describes the underlying sampling and analytical procedures used?

Mortality: The sampling and quality assurance information can be found in Model State Vital Statistics Act and Regulations Revised April 1995, DHHS publication (PHS) 95-1115 (http://www.cdc.gov/nchs/data/misc/mvsact92aacc.pdf). Documentation is also available at http://wonder.cdc.gov/wonder/help/mort.html Table HH2 data source: CDC. CDC WONDER. Compressed Mortality File, Underlying Cause of Death. http://wonder.cdc.gov. The complete web-link pathway from the CDC WONDER Home Page is: à Mortality underlying cause of death à Mortality for 1999 2001 with ICD 10 codes For 2002 data: National Center for Health

Statistics (NCHS). 2004. Deaths Final Data 2002. National Vital Statistics vol 53 no. 5 http://www.cdc.gov/nchs/data/nvsr/nvsr53/nvsr53_05.pdf Regional Mortality: Mortality rates age-adjusted for the 2000 U.S. standard population (rates per 100,000) for the years 1979 through 2001 were compiled through use of CDC WONDER. Maps were created for this same data. See attached file, Calculating data for GIS mortality maps and regional trends.doc , for more information on the calculation of rates. Data file for maps can be found in file MapDataSummaries.xls . Incidence: Each data registry uses its own procedure for collecting and recording data. A complete list of registries and contact information can be found at http://seer.cancer.gov/registries/. Additionally, the North American Association of Central Cancer Registries (http://www.naaccr.org/) helps to guide all state registries to achieve data content and compatibility acceptable for pooling data and improving national estimates. The SEER team is currently developing computer applications to unify cancer registration systems and to analyze and disseminate population-based data. Table HH2 data source: National Cancer Institute (NCI). Surveillance, Epidemiology, and End Results (SEER)*Stat Database: Incidence SEER 9 Regions Public Use, November 2003 Sub (1973-2001), NCI, DCCPS, Surveillance Research Program, Cancer Statistics Branch, released April 15, 2004, based on the November 2003 submission. http://wonder.cdc.gov/seerJ.html. The complete web-link pathway from the CDC WONDER Home Page (http://wonder.cdc.gov/) is: à Cancer Incidence à SEER Cancer Query System à SEER Incidence and US Mortality Statistics Selection criteria: Age Adjusted Rates, Standard Population = 2000 Census, Nine SEER Registry, All Sites; All Ages

T3Q2 Is the complete data set accessible, including metadata, data-dictionaries and embedded definitions or are there confidentiality issues that may limit accessibility to the complete data set?

Mortality: The data can be accessed up to the county level through the electronic data warehouse for CDC at http://wonder.cdc.gov. Individual level data are not available due to confidentiality issues. Incidence: Yes. The public-use data set can be accessed at http://seer.cancer.gov/publicdata/. Data located here includes SEER incidence and population data associated by age, sex, race, year of diagnosis, and geographic areas (including SEER registry and county). Furthermore, SEER distributes tools, such as the SEER*Stat (http://seer.cancer.gov/seerstat/) and SEER*Prep (http://seer.cancer.gov/seerprep/) software, for the analysis of SEER and other cancer databases. SEER data and resources are made available on the sites listed above, free of charge. Individual level data are not available due to confidentiality issues.

T3Q3 Are the descriptions of the study or survey design clear, complete and sufficient to enable the study or survey to be reproduced?

Mortality: Yes. Virtually all deaths from the 50 states, including District of Columbia, submit mortality data to the NVSS at NCHS. The recommended certificate of death is posted at http://www.cdc.gov/nchs/data/dvs/DEATH11-03final-ACC.pdf. The documentation for the mortality data set is http://wonder.cdc.gov/wonder/help/mort.html. Incidence: Yes. See questions T3Q1 and T3Q2 above.

T3Q4 To what extent are the procedures for quality assurance and quality control of the data documented and accessible?

Mortality: See answer to T3Q1 Incidence: The SEER Program is considered as the standard for quality among cancer registries around the world. Quality control has been an integral part of SEER since its inception. Every year, studies are conducted in the SEER areas to evaluate the

quality and completeness of the data being reported (SEER's standard for case ascertainment is 98 percent). In some studies, a sample of cases is reabstracted to evaluate the accuracy of each of the data elements collected from the medical records. In other studies, targeted information gathering is performed to address specific data quality needs. Computer edits also are used by registries to ensure accurate and consistent data.

T4Q1 Have appropriate statistical methods been used to generalize or portray data beyond the time or spatial locations where measurements were made (e.g., statistical survey inference, no generalization is possible)?

Not applicable

T4Q2 Are uncertainty measurements or estimates available for the indicator and/or the underlying data set?

Not applicable

T4Q3 Do the uncertainty and variability impact the conclusions that can be inferred from the data and the utility of the indicator?

Not applicable

T4Q4 Are there limitations, or gaps in the data that may mislead a user about fundamental trends in the indicator over space or time period for which data are available?

Mortality: The mortality data on the Compressed Mortality File at http://wonder.cdc.gov/mortSQL.html are based on records for all deaths occurring in the fifty states and the District of Columbia. Deaths to foreign residents are excluded. Deaths to residents who died abroad are not included on this file. The difference in overall cancer mortality rate for 1999 quoted in ROE03 (201.6 per 100,000) versus the rate quoted in this updated reported (200.8 per 100,000) is that the rate in ROE03 was based on 1990 Census projections and the rate has since been adjusted for 2000 standard population. Cancer mortality rates are based on under-lying cause-of-death as entered on a death certificate by a physician. Some individuals may have had competing causes of death. When more than one cause or condition is entered by the physician, the underlying cause is determined by the sequence of conditions on the certificate, provisions of the ICD, and associated selection rules and modifications. (CDC WONDER database) The 2002 mortality rates were not yet available in CDC WONDER database at the time of this indicator write-up. The age-adjusted mortality rates were obtained through National Center for Health Statistics publication. For purposes of comparison, it should be noted that mortality rates reported by NCHS reports differ slightly from those rates reported by CDC WONDER. NCHS uses U.S. Census Bureau population estimates for all age groups; CDC WONDER uses birth certificate data for the Under 1 Year age group and uses U.S. Census Bureau population estimates for all other age groups. The International Classification of Diseases 9th Revision (ICD 9) codes are used to specify underlying cause of death for years 1979 - 1998. Beginning in 1999, cause of death is specified with the International Classification of Diseases 10th Revision (ICD 10) codes. The two revisions differ substantially, and to prevent confusion about the significance of any specific disease code, data queries are separate.
Regional data: Mortality data are not available for the U.S. territories in CDC WONDER. Thus, Regions 2 and 9 are calculated to include only States. Note that for the years 1979, 1981-1989, and 2001, if the user selects a WONDER query for the United States with data grouped by state, or selects a WONDER query for a specific state, WONDER reports state population figures that

do not add up to the national population reported by WONDER. This is because the two different sets of populations come from different U.S. Census population estimates. (For all other years, these two sets of population data are the same.) Incidence: The SEER Program currently collects and publishes cancer incidence and survival data from 14 population-based cancer registries and three supplemental registries covering approximately 26 percent of the US population. However, the sampled population is representative of the national population.

Data set name(s): Indicator Maps
Cancer Mortality by EPA Region
Chronic Obstructive Pulmonary Disease (COPD) by EPA Region

Data set date(s): 1999-2001

Data source: CDC. WONDER Compressed Mortality 1999–2001 with ICD-10 Codes. http://wonder.cdc.gov/mortICD10J.html.

Data description: Mortality rates are for the U.S. population over the combined years of 1999, 2000, and 2001 (rates per 100,000 population, age-adjusted to the 2000 U.S. standard population). Underlying data were compiled from the WONDER system by age group and by state. Age-adjusted regional mortality rates for the period 1999-2001 were generated using the procedure described in "Calculation of Age-Adjusted Regional Mortality Rates Using State Data from CDC WONDER" (attached).

The mortality rates for each indicator were broken down into equal quartiles

ICD Codes:
Cancer: ICD-10 codes C00 – C97.
COPD: ICD-10 codes J40 – J47.

Supporting data files (represent a subset of the 1979-2001 regional trend data):
MapDataSummary.xls (see attached)

Calculation of Age-Adjusted Regional Mortality Rates Using State Data from CDC WONDER

Mortality rates age-adjusted for the 2000 U.S. standard population (rates per 100,000) for the years 1979 through 2001 were compiled as detailed below. (Mortality data for 2002 and later are not yet available through CDC WONDER.)

Identifying relevant data

- For each mortality indicator (e.g., cancer, cardiovascular disease, asthma) was reviewed based on EPA's 2003 Draft Report on the Environment and the National Center for Health Statistics (NCHS) report that discusses ICD (International Classification of Diseases) to obtain the ICD codes that describe each of these indicators. Due to a revision of the ICD system in 1999, we obtained ICD-9 codes for the years 1979-1998 and ICD-10 codes for the years 1999-2001.

Downloading and organizing the data

- Mortality data was accessed through CDC's WONDER database (http://wonder.cdc.gov/).

- We downloaded a file for each year that regional mortality indicators were requested (1979-2001) and saved these data into separate sheets in Microsoft Excel (one workbook for each indicator). Both raw and compiled data are presented in each of these Excel workbooks. For example, '1999D' is the sheet with the raw data for the year 1999 and '1999' is the sheet with the calculations for that year.

- The calculation worksheet is organized as follows:

    - The first table arranges the raw data by state and age group. For each age group and state, two numbers are presented: the mortality cases and the associated population.

    - The second table (Regional Breakdown) groups the mortality cases and populations for the 50 states plus Washington, D.C., into the ten EPA regions, to obtain the total cases and total population for each region, by age group. (For a map of the EPA regions, refer to http://www.epa.gov/epahome/whereyou-live.htm.)

    - The third table (Regional Breakdown [combine certain groups]) merges two sets of age groups with each other (5-9 years with 10-14 years; 15-19 years with 20-24 years) to match the age-adjustment methodology used to calculate age-adjusted mortality rates in CDC WONDER (see below).

    - The fourth table on the worksheet (Regional Breakdown [Computation of Age-Adjusted Rate Components]) presents the basic steps of calculating age-adjusted mortality rates for each region (see below).

    - The final table on the Excel sheet (Regional Summary) lists the age-adjusted rates for each region and for the entire US; these values are compiled into the summary sheet

that covers all regions for all years for an indicator, accompanied by a trend chart that graphically depicts the regional data.

Calculating age-adjusted regional rates

The steps followed in calculating age-adjusted regional rates are detailed below, followed by a sample calculation.

Step 1

Using the following equation, the *crude death (or mortality) rate* is obtained by dividing the mortality cases ($Cases_i$) by the population for that age group ($Population_i$), then by multiplying by 100,000 (to get the cases per 100,000). This is done for each age group within each region. (The "i" subscript is included to indicate that this calculation is performed for several age groups, i.e. $CrudeRate_{<1\ year}$, $CrudeRate_{1\text{-}4\ years}$, etc.)

$$CrudeRate_i = \frac{Cases_i}{Population_i} \times 100{,}000$$

Step 2

For each age group, the weighted age-adjusted factor is calculated, using 2000 U.S. standard population factors provided by NCHS (See Table 1 below). The age-specific crude death rate is multiplied by that age group's standard population for the year 2000 ($2000\ Population_i$), and then divided by the total standard population for the year 2000 ($2000\ Population_{total}$). (For details on the standard population, see "Age-Adjustment of Death Rates" on the web page http://wonder.cdc.gov/wonder/help/mort.html[1]).

$$WeightedFactor = CrudeRate \times \frac{2000\,Population_i}{2000\,Population_{total}}$$

Table 1. United States Standard Population*

Age	Number
Under I year	13,818
1-4 years	55,317
5-14 years	145,565
15-24 years	138,646
25-34 years	135,573
35-44 years	162,613
45-54 years	134,834
55-64 years	87,247
65-74 years	66,037
75-84 years	44,842
85 years and over	15,508
All ages	1,000,000

* Based on year 2000 projected population

[1] The source cited by CDC WONDER for the age-adjustment data is the following NCHS report: *Anderson RN, Rosenberg HM. Age standardization of death rates: Implementation of the year 2000 standard. National Vital Statistics Reports; Vol 47 No 3. Hyattsville, Maryland. National Center for Health Statistics. 1998.*

Step 3

The age-adjusted mortality rate is then obtained by adding together the individual weighted factors for each age group:

$$AgeAdjustedRate = \sum_{i} WeightedFactor_i$$

Sample Calculation

Table 2 (below) shows the output of a sample calculation of the age-adjusted mortality rate for cancer (across all age groups) in EPA Region 9 in 1982. (Note that we have presented this in a format that is not used in the Excel workbook, in order to illustrate the calculations for a single region. However, all of the steps and calculations are identical.)

For each age group, the number of mortality cases is the sum of the mortality cases for Arizona, California, Hawaii, and Nevada, for that age group in that year; the population is the sum of the populations of these same four states for that age group in that year. For example, for the 35-44 years age group, the number of cancer mortalities for these states are 112 (Arizona), 1,413 (California), 53 (Hawaii), and 55 (Nevada), with the sum equaling 1,633. Similarly, the population is the sum of the respective states 338,654 (Arizona), 3,152,885 (California), 124,743 (Hawaii), and 120,463 (Nevada), which equals 3,736,745. The crude rate (43.70) for this age group is therefore equal to the total number of mortality cases (1,633) divided by the total population (3,736,745), multiplied by 100,000.

Table 1. Cancer Mortality, All Age Groups, EPA Region 9, 1982.

Age Group (Years)	Mortality Cases	Year 1982 Population	Crude Death Rate	2000 Std. Population	Weighted Factor
<1 year	32	515,809	6.20	13,818	0.09
1- 4 years	108	1,824,635	5.92	55,317	0.33
5- 9 years	114	1,977,487	N/A		
10-14 years	94	2,188,828	N/A		
5-14 years	208	4,166,315	4.99	145,565	0.73
15-19 years	139	2,432,939	N/A		
20-24 years	180	2,884,175	N/A		
15-24 years	319	5,317,114	6.00	138,646	0.83
25-34 years	694	5,551,792	12.50	135,573	1.69
35-44 years	1,633	3,736,745	43.70	162,613	7.11
45-54 years	4,662	2,793,603	166.88	134,834	22.50
55-64 years	11,283	2,679,802	421.04	87,247	36.73
65-74 years	15,423	1,876,606	821.86	66,037	54.27
75-84 years	11,397	912,865	1248.49	44,842	55.98
85+ years	4,424	269,593	1640.99	15,508	25.45
Unknown	27	0	N/A	0	0
Total	*50,210*	*29,644,879*	*N/A*	*1,000,000*	N/A
Age-Adjusted Mortality Rate					**205.7**

As described above, the Weighted Factor is the Crude Death Rate multiplied by the 2000 Standard Population for that age group and divided by the total Standard Population (the total of the age-group populations). For example, the weighted factor for the 35-44 years age group is the crude rate (43.70) times the 2000 population for that group (162,613), divided by the total 2000 Standard Population (1,000,000), which equals 7.11. The 1982 Region 9 age-adjusted mortality rate, 205.7, is the sum of the weighted factors of all age groups.

Notes:

- For the 5-14 and 15-24 years categories it is necessary to merge two sets of age ranges to match the age-adjustment grouping used within WONDER.

- For mortality indicators that were also compiled for children (ages 0-19 years), we only used the data for the age groups 0-1, 1-4, 5-9, 10-14, and 15-19 years, and then we age-adjusted these data using a set of age-adjustment factors that only cover to age 19 years.

- For the file that compiles birth defect mortality rates, the only data used from CDC WONDER are for the <1 year age group, so the crude rate equals the age-adjusted rate.

- Although data was queried for individual states to compile regional data, we did this by querying data for the entire United States from the CDC WONDER system, and specifying that the data be grouped by age and by year. Due to a quirk of CDC WONDER, if the user selects a single state for a query (instead of the entire United States), the population data are taken from a different data source: there are small discrepancies between these numbers and so the state-specific query should not be used to verify the compiled data.